**Fantasy Premier League. Выбор команды перед первым туром**

В [первой заметке](https://baguzin.ru/wp/?p=25620) цикла я рассказал, что ряд менеджеров Fantasy Premier League демонстрирует высокие результаты на протяжении многих сезонов. Я выделил ТОР-10k менеджеров, показавших наивысшие результаты в сезонах 2017/2018–2021/22 (*элиту*), а также группу сравнения – аккаунты, отобранные случайным образом (*поляну*). Во [второй заметке](https://baguzin.ru/wp/?p=25833) я показал, какие возможности для извлечения данных с сайта Fantasy Premier League предоставляет API и Excel Power Query. Начиная с этой заметки я буду рассказывать о различных аспектах игры, сравнивая поведение элиты и поляны. А начну с выбора команды к GW1.

### Вратари

3 августа 2022 г. основной вратарь Лестера Каспер Шмейхель неожиданно перешел в Ниццу. Дэнни Уорд занял его место, а стоил он всего £4М. Многие прельстились уменьшить инвестиции во вратарей, и взяли Уорда и второго вратаря Лестера – Даниэля Иверсена.

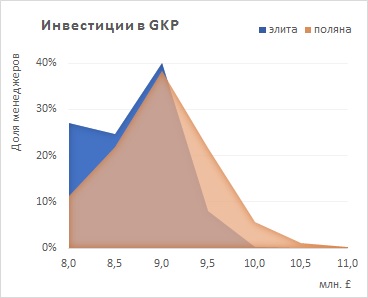


Рис. 1. Распределение инвестиций во вратарей

Хотя распределения заметно отличаются, средние значения довольно близки: элита – £8,7М, поляна – £9,0М. В приложении показано, что отличия двух распределений статистически значимы.

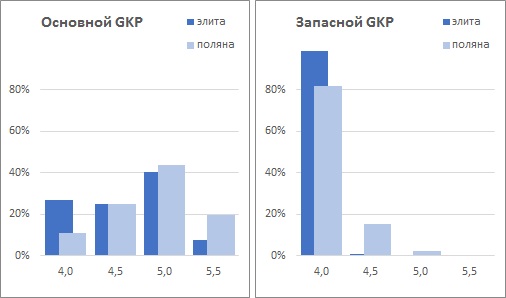


Рис. 2. Распределение инвестиций раздельно в основного и запасного вратарей

Видно, что по основным вратарям за £4,5М и £5М значения близкие. Отличие – в предпочтении элитой бюджетного варианта за £4М над премиальным за £5,5М. 9873 менеджера элиты (из 10 000) запасным взяли вратаря за £4М. Среди поляны таких было меньше – 8195. Если сравнивать пары вратарей, то самый популярный вариант в обеих группах – £5М + £4М. А вот в более бюджетных вариантах – 4 + 4 и 4,5 + 4 – доля элиты выше.

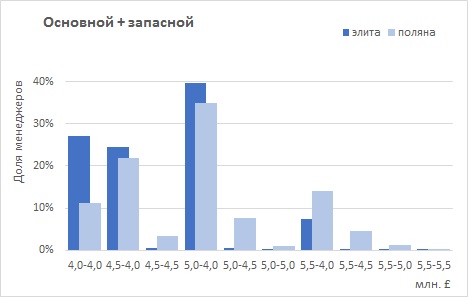


Рис. 3. Предпочтения по паре вратарей

Элита более концентрирована в своем выборе: всего 42 различных вратаря. А на десять самых популярных вратарей пришлось 98% загрузок. Представители поляны остановили свой выбор на 56 вратарях. На десять самых популярных пришлось 86% загрузок:

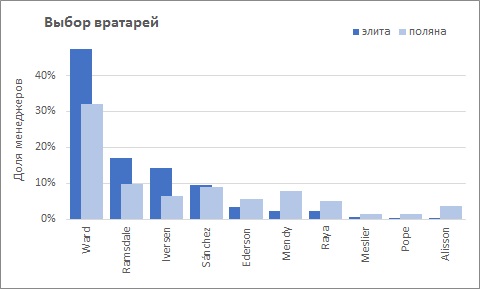


Рис. 4. Предпочтения по загрузкам вратарей

### Деньги в банке

В среднем элита оставляет в банке £0,26М, поляна – £0,23М. Элита более склонна иметь резерв в £0,5М:

Рис. 5. Распределение резерва в банке

Удивительно, но каждый десятый менеджер оставил в банке £1М и более! Вновь показано, что отличия являются статистически значимыми (см. Excel-файл, лист Bank).

### На скамейке

В среднем элита оставляет на скамейке £17,5М, поляна – £17,9М. Показано, что отличия являются статистически значимыми (см. Excel-файл, лист Bench).

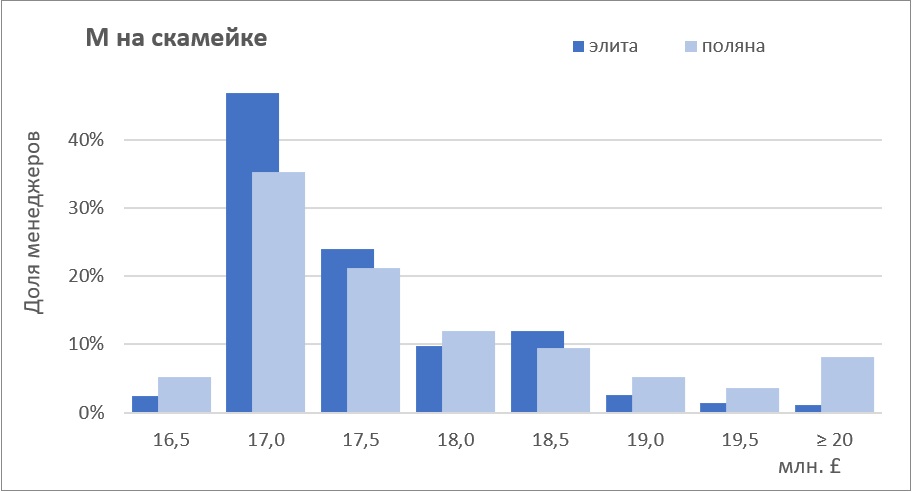


Рис. 6. Распределение инвестиций в скамейку

Поляна более склонна к крайностям: оставить на скамейке слишком мало или слишком много. £16,5М можно оставить на скамейке лишь в варианте с двумя защитниками за £4М, а как мы увидим далее, элита начала сезон с более мощной защитой.

Некоторые игроки элиты оставили на скамейке £25М и более. Изучив внимательнее, я обнаружил, что такие менеджеры заиграли в первом туре Bench Boost. А для прикола многие менеджеры в случае BB любят оставлять на скамейке более дорогих игроков))

### Байес

Прежде, чем продолжить исследование, настал удобный момент сказать несколько слов о том, как в принципе элитные менеджеры делают свои прогнозы. Их высокие результаты являются следствием многих факторов, но один из решающих – [мышление в стиле Байеса](http://baguzin.ru/wp/?p=6355). Я не думаю, что все элитные менеджеры знакомы с теорией Байеса, но интуитивно они ей следуют. Суть относительно проста. У нас есть априорная вероятность – исходная статистика. У нас есть вероятностный опыт, т.е. событие, которое может закончиться так или иначе. На основании априорной вероятности и того, каким результатом завершился опыт, возникает апостериорная вероятность – лучшее предположение о будущем.

Можно сформулировать немного иначе. Текущие убеждения после поступления информации превращаются в новые убеждения. Информация может поддерживать текущие убеждения, а может не соответствовать им. Но в любом случае новые убеждения будут базироваться как на текущих убеждениях, так и на поступающей информации.

Допустим, у нас есть обширная статистика по защитнику Ливерпуля Trent Alexander-Arnold: результаты за последние сезоны, выходы в старте, очки дома и на выезде, против сильных и слабых команд и т.п. В первом матче сезона 2022/23 Ливерпуль на выезде встречается с новичком Премьер-лиги – Фулхэмом. На основе статистики можно ожидать, что ТАА с вероятностью 65% совершит как минимум одно действие. В байесовской терминологии эта вероятность называется априорной. Соотношение вероятностей отражено на верхнем бруске. Левая область «Без действия» составляет 35%, правая «С действием» – 65%.

Изображение выглядит как стол

Автоматически созданное описание

Рис. 7. Новая информация не изменила убеждения

Перед матчем FUL-LIV мы не знаем, совершит ли ТАА действие. Результат ТАА в матче – это наш вероятностный опыт. Матч закончился со счетом 2:2, а ТАА набрал лишь очко. Небольшая стрелка показывает, что результат этого события почти не изменил вероятности при переходе от верхнего бруска к нижнему.

Нижний брусок отражает лучшее суждение, чего ожидать от ТАА в будущем. Это наша апостериорная вероятность, появившаяся путем взаимодействия априорной вероятности и конкретного результата опыта. На этом бруске области «С действием» и «Без действия» почти не изменились. Новость о том, что ТАА в одном матче набрал лишь очко почти не повлияла на наши ожидания.

И, действительно, как мы увидим далее, элита не отгрузила ТАА после плохого первого тура. Разовая неудача не опрокинула прогноз, а лишь немного его скорректировала. Что делают многие новички fpl? Преувеличенно реагируют на одиночный опыт. Они придают единичному результату слишком большой вес, значительно уходя от априорной вероятности. Что делает элита fpl? Учитывает единичный опыт, но придает ему соразмерный вес. Априорная вероятность не отметается, а корректируется в апостериорную.

В нашем примере потребовалось шесть туров (лишь один из которых был удачным), чтобы более половины менеджеров элиты отгрузили ТАА:

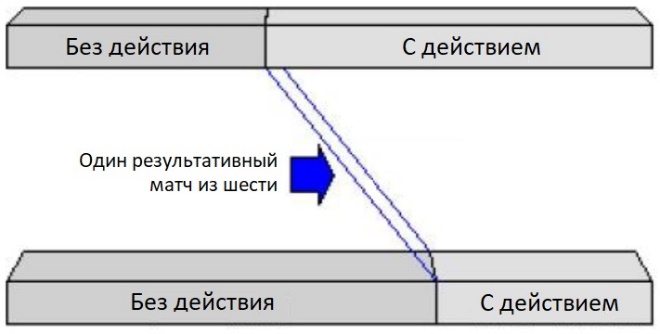


Рис. 8. Более значительная новая информация изменила убеждения

Теперь наш опыт заключается в том, что только в одном туре из шести ТАА совершил действие. Влияние стрелки значительно: смещение при переходе от априорной вероятности к апостериорной существенно.

### Расстановки

Теперь я могу пояснить, почему на старте сезона 2022/23 элита предпочла расстановки с пятью и четырьмя защитниками. Всё дело в том, что по итогам предыдущего сезона защитники были столь успешны, что не взять их, представлялось ошибкой:



Рис. 9. Высокие позиции защитников по итогам сезона 2021/22

Часть поляны действовала на основе устаревшей парадигмы: защитники набирают мало, а нападающие много, поэтому следует предпочесть расклады с тремя защитниками. Различия в предпочтении элиты и поляны хорошо видны на следующей диаграмме:

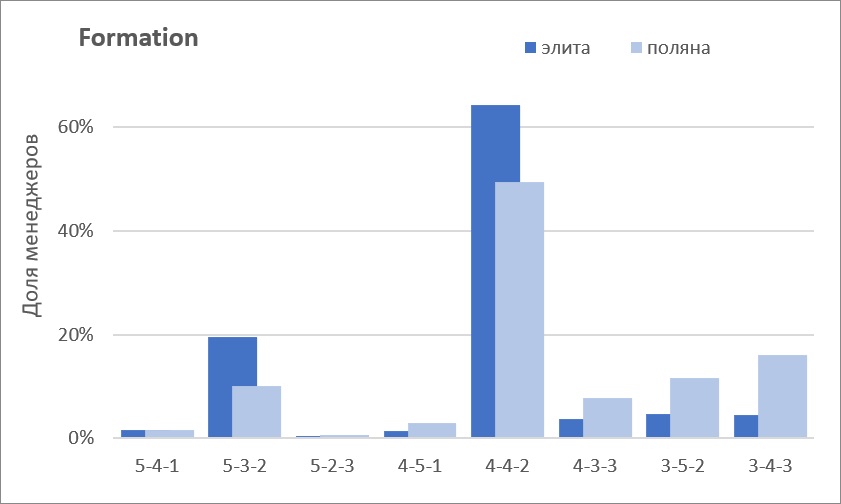


Рис. 10. Расстановки

Важно понимать, что сезон 2022/23 разворачивается не так как предыдущий. Элитные защитники не показывают достойных результатов. Наиболее популярной в середине сезона является расстановка 3-4-3. Весьма вероятно, что на старте следующего сезона элита не будет выбирать расстановки с 5 защитниками!

### Бюджет по линиям

Различия также можно обнаружить в распределении бюджета по линиям. Элита более консолидирована, т.е. выбирает меньшее число игроков, поэтому ее пики выше. Пусть вас не смущает небольшое отличие в средних значениях – распределения отличаются существенно.

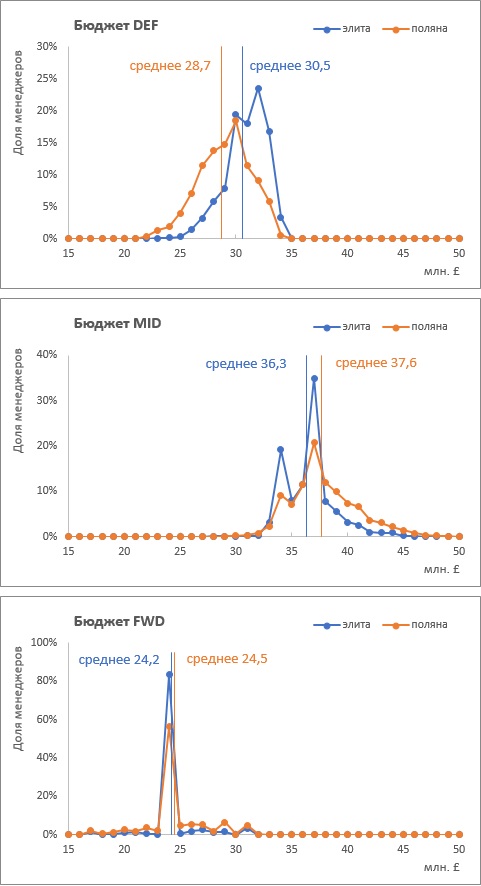


Рис. 11. Распределение бюджета по линиям

### Выбор игроков

Распределения демонстрируют б***о***льшую концентрацию мнений у элиты по сравнению с поляной. Это говорит о том, что в элите меньше различий в суждениях. Решения принимаются на основе приблизительно одних и тех же факторов. Суждения поляны разнообразнее, на решения влияют различные факторы. Например, выбор игроков из фанатской команды.

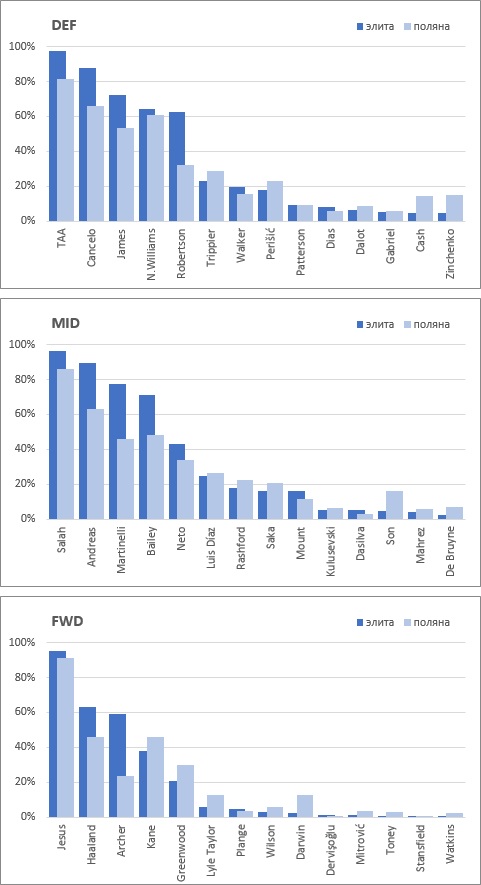


Рис. 12. Выбор игроков по линиям

### Выбор капитана

… также более консолидирован у элиты:

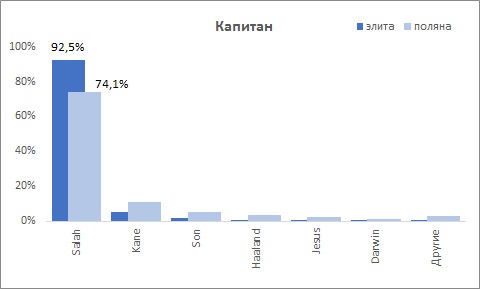


Рис. 13. Выбор капитана к GW1

Поляна чаще выбирает капитана среди дифференциалов.

### Баланс команды

Большинство менеджеров начали сезон с двумя премиальными игроками:



Рис. 14. Число премиальных игроков

Но и здесь, мнение элиты более консолидировано: только 859 менеджеров взяли иное число премиальных игроков. Среди менеджеров поляны таких оказалось 2188 или в 2,5 раза больше.

Аналогичная ситуация и с довесочками. К таковым относят голкиперов и защитников за £4М, полузащитников и напов за 4,5. Напомню, у вратарей играющий Дэнни Уорд формально является довесочком. Статистика вратарей рассмотрена ранее отдельно. Поэтому здесь речь о числе довесочков в трех других линиях.

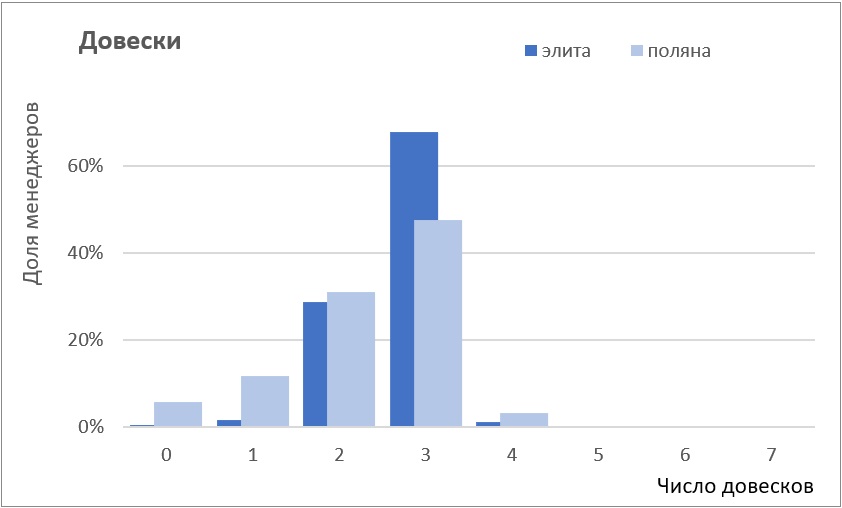


Рис. 15. Число довесочков среди DEF, MID и FWD

Видно, что элита берет трех или двух довесочков. Чуть менее 20% поляны не готовы терпеть столь слабую скамейку и ограничиваются 0–1 довеском. Но некоторые менеджеры поляны, наоборот, пускаются во все тяжкие и берут 4–7 довесков. Вот пример такого экстремизма (аккаунт 1877051)

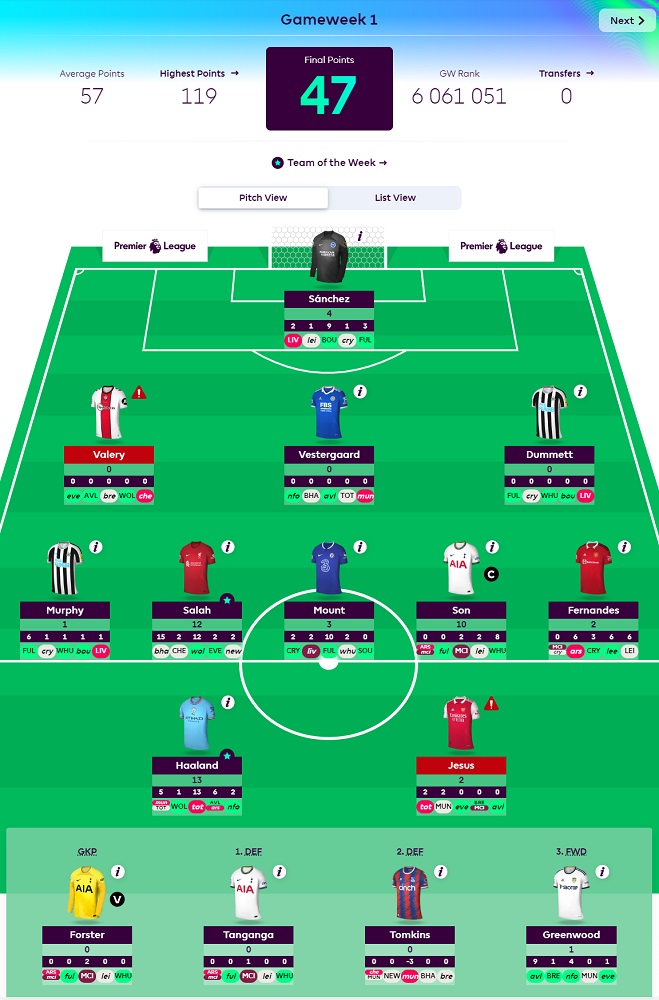


Рис. 16. Пример максимально разбалансированной команды: 8 довесков, в т.ч. GKP

Как правило, довески не появляются в старте команд Премьер-лиги, но бывают исключения. Возникают они, когда игрок, получивший минимальный ценник при открытии сезона FPL, и прочно сидящий в своей команде на скамейке, переходит в новую команду и получает перспективы стать основным.

Сезон 2022/23 является нетипичным. До старта GW1 несколько игроков перешли из топовых команд в более слабые. Все эти игроки были по достоинству оценены менеджерами FPL. Такая ситуация позволила сформировать экстремально бюджетную и при этом играющую скамейку.

Лучший запасной полузащитник – Andreas Pereira. Уже после объявления его цены в £4,5М 11 июля 2022 г. перешел из МЮ в Фулхэм. Мало того, что он дёшев, так он еще довольно высоко играет, и является вторым штатным пенальтистом!

Второй пример – защитник Neco Williams. Он также после объявления цены в £4М 10 июля 2022 г. перешел из Ливерпуля в Ноттингем. В первой части сезона он постоянно появлялся в старте и даже заработал три клиншита. После окончания перерыва на Кубок мира он потерял место в основе.

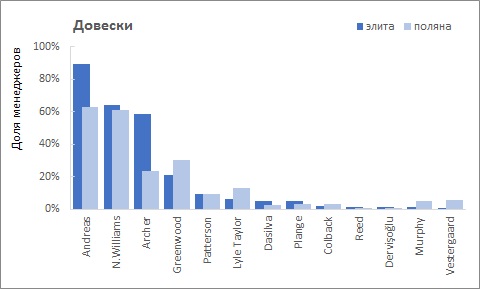


Рис. 17. Наиболее популярные довески

### Баланс внутри линий

Что можно заметить на рис. 18?

**Защита.** Элита в начале сезона делала ставку на премиальных игроков, поэтому владение в диапазоне £6–7,5М у элиты существенно выше. 3,3 игрока против 2,4 у поляны. Владение довесками за £4М близко у обеих групп. А вот промежуточных игроков за £4,5–5,5М элита загрузила намного меньше, чем поляна: 1,0 против 1,7.

**Полузащита.** Элита более бескомпромиссна. В нижнем ценовом диапазоне £4,5–6М + Салах средняя загрузка = 4 против 3,3 у поляны. И наоборот, в среднем ценовом диапазоне £6,5–12М средняя загрузка у элиты = 1,0, у поляны = 1,7.

**Нападение.** В этом сезоне произошла фантастическая поляризация. Такое впечатление, что были предложены нападающие лишь в трех ценовых диапазонах: 4,5, 8,0 и 11,5. Но и здесь поляризация элиты выше. В этих трех диапазонах средняя загрузка у элиты составила 2,9 против 2,5 у поляны.

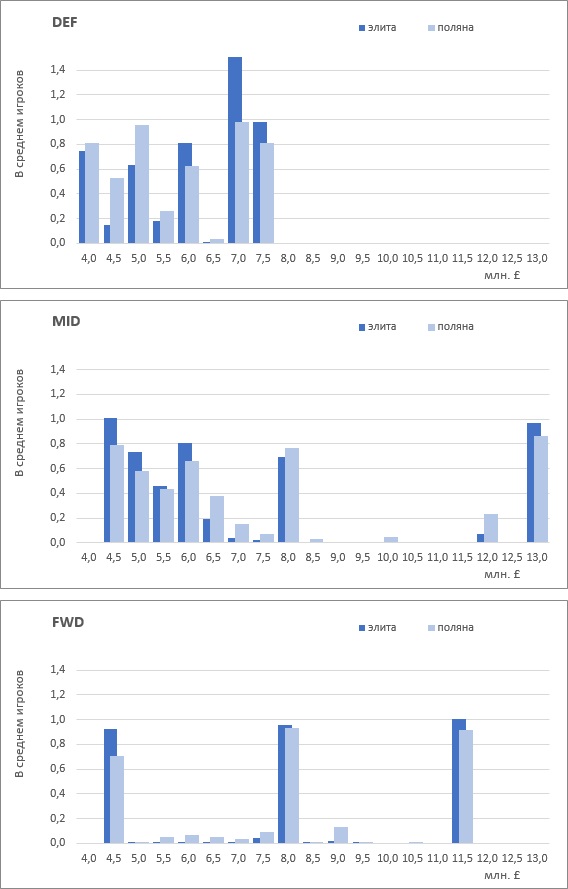


Рис. 18. Распределение бюджетов внутри линий

### Команды игроков

В выборе команд игроков основного состава элита, как и в остальных аспектах, характеризуется большей поляризацией. Из LIV, ARS и MCI элита взяла в среднем 7,1 игрока, а поляна – 6,2. С другой стороны, из наименее популярных 10 команд элита взяла в среднем лишь 0,14 игрока, а поляна – 1,3.

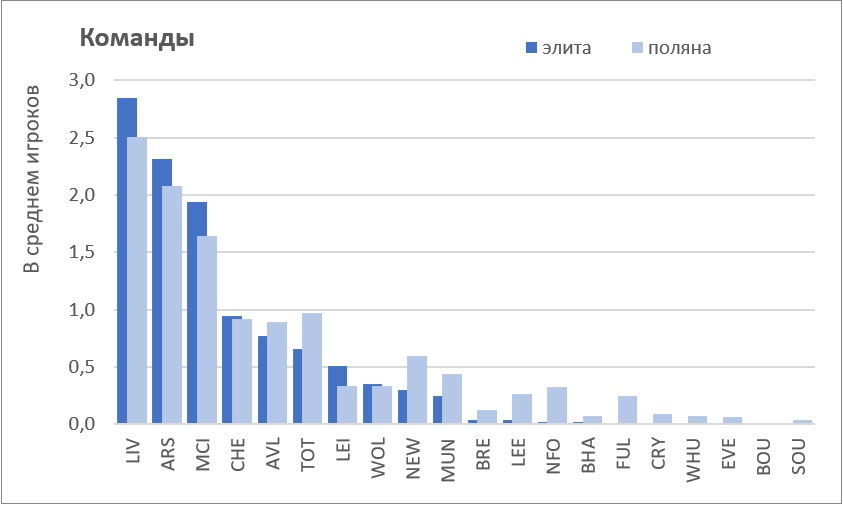


Рис. 19. Выбор команд игроков основного состава

### Число игроков из фанатской команды

Строго говоря, для корректного сравнения, нужно чтобы распределение фанатских предпочтений среди элиты и поляны совпадали:

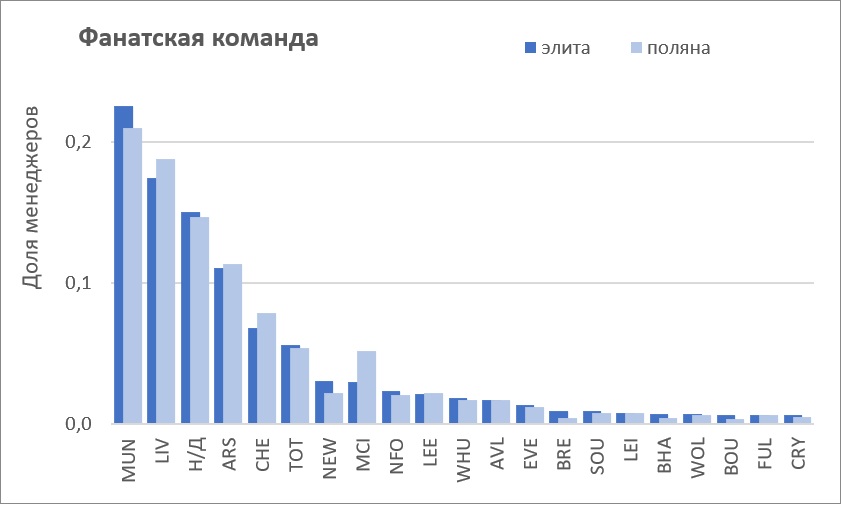


Рис. 20. Фанатские предпочтения

Реальное совпадение не слишком хорошее, поэтому следующую диаграмму я представляю с некоторыми оговорками.

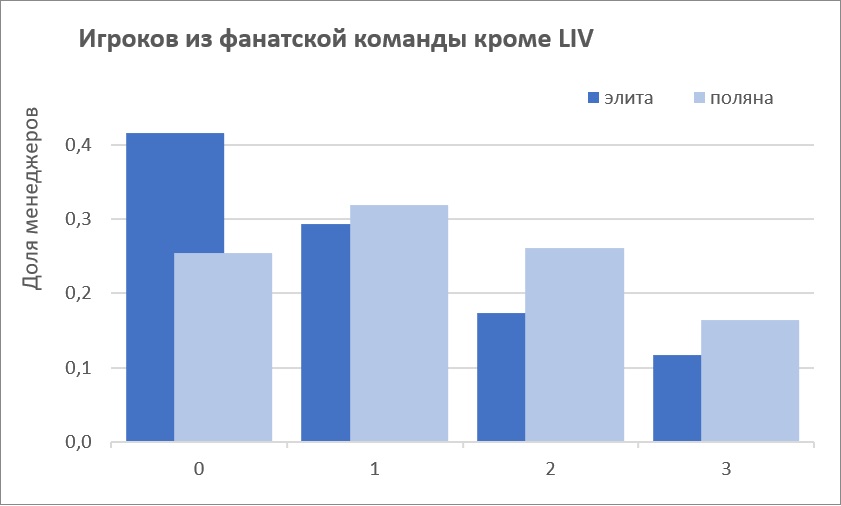


Рис. 21. Количество игроков из фанатской команды

Здесь я исключил фанов Ливерпуля из обоих групп, поскольку 90% из них взяли трех игроков LIV, что хорошо совпадает со средним выбором игроков Ливерпуля, равным 2,85 (см. рис. 19).

Видно, что элита меньше подвержена влиянию фанатских искажений. В среднем элитный менеджер из фанатской команды берет 1 игрока, а менеджер поляны – 1,3. Распределение поляны сильно смещено вправо относительно распределения элиты.

### Результаты по итогам GW19

Понятно, что успехи элиты определяются не только выбором команды к GW1. К тому же сезон 2022/23 не похож на предыдущий. И всё же любопытно оценить, результаты элиты и поляны на момент публикации заметки:

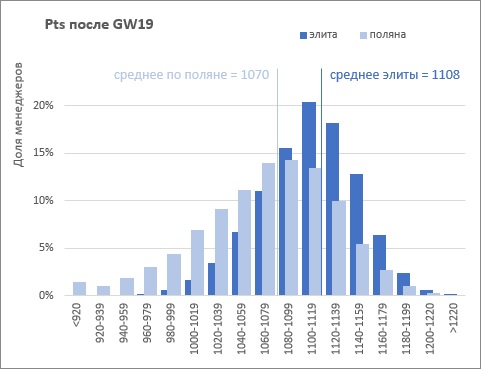


Рис. 22. Распределение результатов элиты и поляны по окончанию GW19

В среднем элита за 19 туров обыграла поляну на 38 очков, т.е. за тур элита набирает на 2 Pts больше. На мой взгляд, это немало))

### Основные выводы

1. Суждения элиты по любому вопросу более единодушны. Поляна демонстрирует больше различий во мнениях.
2. Элита в качестве основного чаще выбирает бюджетного вратаря; в запас всегда берет за £4М.
3. 35% менеджеров элиты перед GW1 оставляют что-то в банке.
4. Скамейка элиты более бюджетная; в половине случаев на скамейке £16,5–17М (£16,5М – теоретически возможный минимум).
5. Элита выносит суждения в стиле Байеса: прогноз основан на балансе долгосрочной статистики и последних результатов.
6. Учитывая успехи защиты в предыдущем сезоне, 84% менеджеров элиты выбрали расстановки 5-3-2 и 4-4-2.
7. 91% менеджеров элиты включили в состав двух премиальных игроков.
8. Элита на 30% меньше подвержена влиянию фанатских искажений.

### Приложение. Извлечение данных из нескольких API-страниц

[Ранее](https://baguzin.ru/wp/?p=25833) я показал, какие возможности для извлечения данных с сайта Fantasy Premier League предоставляет API и Excel Power Query. Чтобы в одном запросе извлечь все данные по элите, я сначала, извлек данные по одному менеджеру, полностью их обработал, затем превратил запрос в функцию, и, наконец, использовал функцию для списка API-страниц. Кратко разберем эти шаги.

#### Извлечение и обработка данных по одному менеджеру

В Excel пройдите *Данные* –> *Получить и преобразовать данные* –> *Из Интернета*. В открывшемся окне вставьте адрес API-страницы <https://fantasy.premierleague.com/api/entry/728021/event/1/picks/> с составом команды лучшего менеджера последних сезонов Fábio Borges. Нажмите Ok. В окне редактора Power Query кликните правой кнопкой мыши на список (List) записи *picks* и выберите *Детализация*:

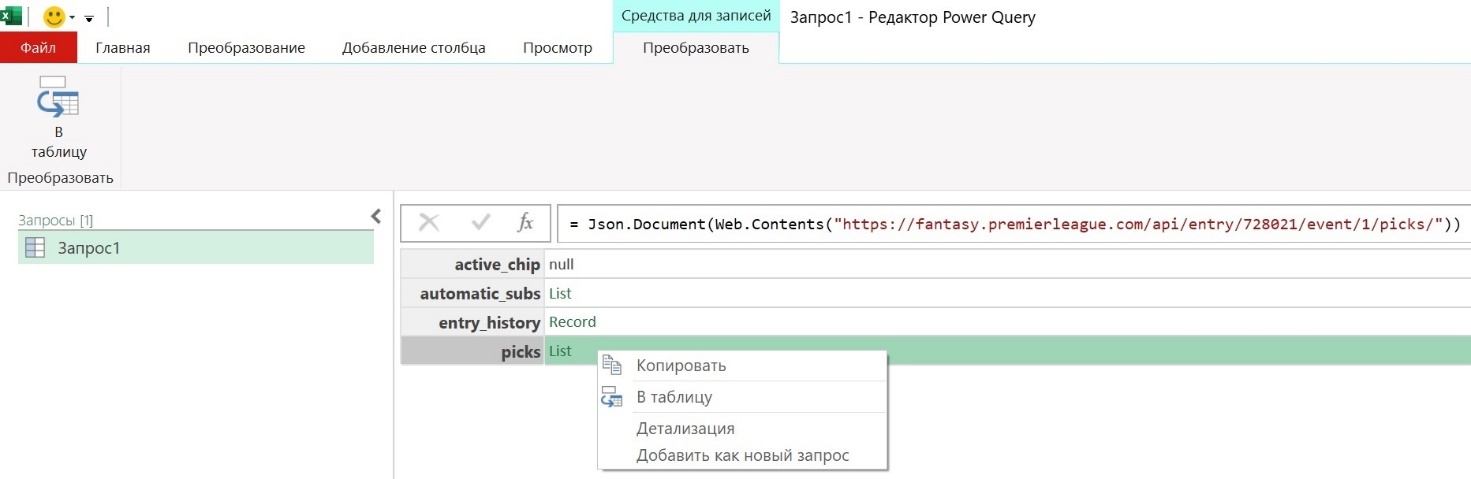


Рис. П1. Выделение списка *picks*

Список раскроется, и вы увидите, что он включает 15 записей (рис. П2). Кликните на кнопке *В таблицу* (1). В окне *В таблицу* выберите разделитель *Пользовательский* (2), оставьте поле разделителя пустым (3).

Изображение выглядит как стол

Автоматически созданное описание

Рис. П2. Преобразование списка записей в таблицу

Список из 15 записей преобразуется в таблицу с одним столбцом *Column1* из 15 строк (рис. П3). Кликните справа от названия столбца по кнопке с двунаправленной стрелкой (1). В открывшемся окне снимите галку с опции *Использовать исходное имя столбца как префикс* (2). Нажмите Ok.

Изображение выглядит как текст

Автоматически созданное описание

Рис. П3. Раскрытие столбца *Column1*

Переименуйте запрос в *element\_pick\_1*:

Изображение выглядит как стол

Автоматически созданное описание

Рис. П4. Запрос *element\_pick\_1*

Создайте запросы *teams*, *element\_types* и *elements\_ID,* как описано в заметке [Извлечение больших данных Fantasy Premier League с помощью API и Excel Power Query](https://baguzin.ru/wp/?p=25833). Все запросы можно найти в приложенном файле Excel. В таблице *elements\_ID* замените id команд и позиций на их имена, используя объединение с запросами *teams* и *element\_types*. Должно получиться что-то типа:

Изображение выглядит как стол

Автоматически созданное описание

Рис. П5. Запрос *elements\_ID*

Объедините запросы *element\_pick\_1* и *elements\_ID*:

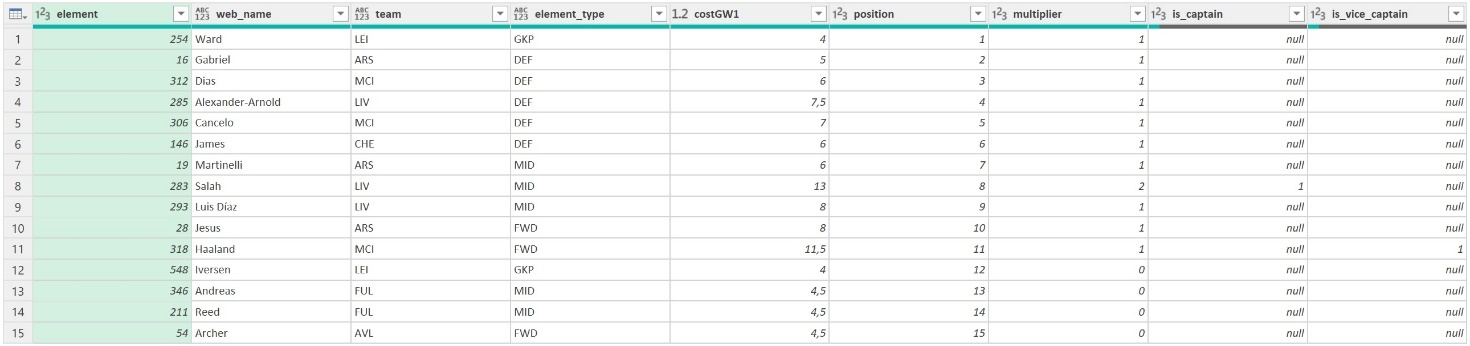


Рис. П6. Запрос *element\_pick\_1* после объединения с *elements\_ID*

#### Превращение запроса в функцию

Превратим запрос *element\_pick\_1* в функцию. Для этого в редакторе Power Query кликните на запросе *element\_pick\_1* и пройдите по меню *Главная* –> *Запрос* –> *Расширенный редактор*:

Изображение выглядит как текст, стол

Автоматически созданное описание

Рис. П7. Запрос *element\_pick\_1* в окне расширенного редактора

Перед строчкой *let* вставьте новую строку

*(Link) as table =>*

Строку…

*Источник = Json.Document(Web.Contents("https://fantasy.premierleague.com/api/entry/728021/event/1/picks/")),*

… замените на…

*Источник = Json.Document(Web.Contents(Link)),*

Запятая в конце строки обязательна. Получится:

Изображение выглядит как текст

Автоматически созданное описание

Рис. П7. Изменение кода для превращения запроса в функцию

Нажмите Ok. Запрос превратится в функцию. Обратите внимание, что значок таблицы перед именем запроса превратился в значок функции. Все ПРИМЕНЕННЫЕ ШАГИ схлопнулись в один. Код функции доступен для редактирования только в расширенном редакторе.

Изображение выглядит как текст

Автоматически созданное описание

Рис. П8. Интерфейс функции *element\_pick\_1*

#### Использование функции для списка API-страниц

Я взял список аккаунтов, входящих в [элиту](https://baguzin.ru/wp/wp-content/uploads/2022/10/Masterstvo-ili-udacha.-Vyborki.xlsx), и на базе этого списка подготовил умную таблицу в Excel. Строки из второго столбца – *Ссылка* – я использую, в качестве адресов, которые по очереди буду «вскармливать» только что созданной функции *element\_pick\_1*.

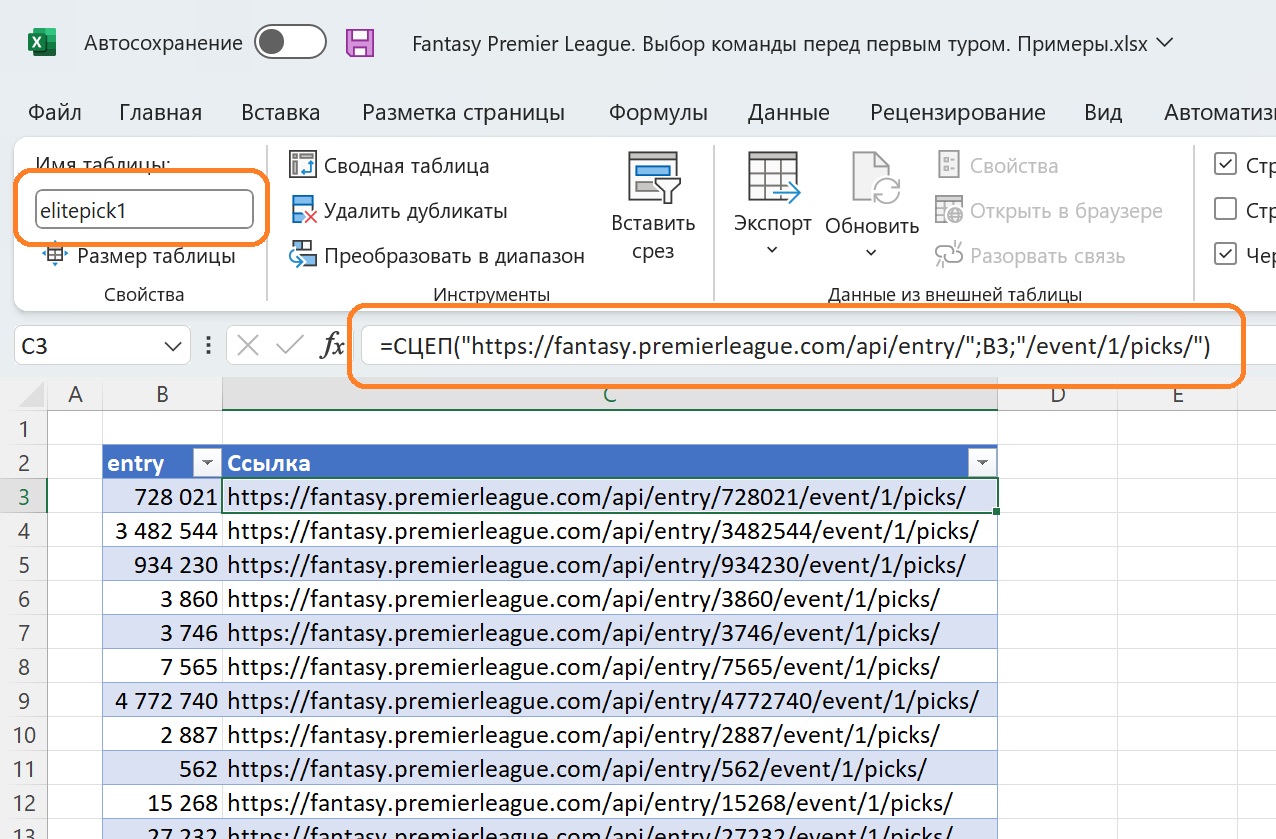


Рис. П9. Умная таблица с адресами API-страниц 10 000 аккаунтов элиты

В Excel кликните на любую ячейку таблицы *elitepick1* и пройдите по меню *Данные* –> *Из таблицы/диапазона*. Откроется редактор Power Query, будет создан запрос *elitepick1*. В редакторе Power Query пройдите *Добавление столбца* –> *Вызвать настраиваемую функцию*. Настройте параметры окна *Вызвать настраиваемую функцию* как указано на рис. П10:

Изображение выглядит как текст

Автоматически созданное описание

Рис. П10. Настройка нового столбца с вызываемой функцией

Появится новый столбец *element\_pick\_1*, содержащий таблицу в каждой строке. Удалите, ставший ненужным столбец *Ссылка*, и раскройте столбец *element\_pick\_1*. Еще раз задайте числовые форматы столбцов – форматы из функции *element\_pick\_1* не переносятся в запрос *elitepick1*.

К сожалению, при попытке загрузить запрос в модель данных, где-то на 200-й строке произойдет сбой. Он связан с тем, что один из запросов не нашел соответствующую API-страницу. Такое может случиться по массе причин. В моем случае в элитный список по формальным признакам «прокрались» несколько аккаунтов, которые в сезоне 2022/23 стартовали не с первого тура. Может быть, их следовало исключить из элиты!? Не знаю… Я оставил, и добавил в код функции *element\_pick\_1* проверку:

Изображение выглядит как текст

Автоматически созданное описание

Рис. П11. Код функции *element\_pick\_1* с проверкой наличия API-страницы

Изменения кода затронули несколько первых строк. Теперь, если аккаунт не создал команду к GW1, запрос вернет пустую строку и продолжит свою работу. 10 000 аккаунтов по 15 игроков, итого 150 000 строк. Скорость загрузки зависит от множества факторов. У меня загрузка заняла около 20 минут.

#### Корректировка списков элиты и сравнения

В материалах [первой заметки](https://baguzin.ru/wp/?p=25620) я выделил ТОР-10k менеджеров, показавших наивысшие результаты в сезонах 2017/2018 – 2021/22 (*элиту*), а также группу сравнения – аккаунты, отобранные случайным образом (*поляну*). После извлечения составов команд элиты к GW1 оказалось, что некоторые команды странные. Например, Jonathan Peel #22746 имел довольно высокий рейтинг по итогам последних пяти сезонов…

Изображение выглядит как стол

Автоматически созданное описание

Рис. П12. История выступлений Jonathan Peel #22746

…но к сезону 2022/23 взял в свою команду двух элитных вратарей за 5,5М. На этом странности не закончились: в текущем сезоне не было сделано ни одного трансфера!

К сожалению, это свойство больших данных: они содержат некоторое количество мусора. Поэтому я удалил из элиты и поляны аккаунты, не сформировавшие команды к GW1, и сделавшие подозрительно мало трансферов к перерыву на мировой чемпионат. Также для ускорения извлечения данных уменьшил размеры поляны до 10k аккаунтов.

Полные списки аккаунтов элиты и поляны приведены в файле Excel.

#### Проверка статистической значимости различий

На рис. 1 приведены два распределения инвестиций во вратарей: элиты и поляны. Являются ли отличия статистически значимыми? Ответ можно получить с помощью критерия согласия χ2. Технику проверки см. по [ссылке](http://baguzin.ru/wp/?p=6056) и в приложенном файле Excel на листе GKP.

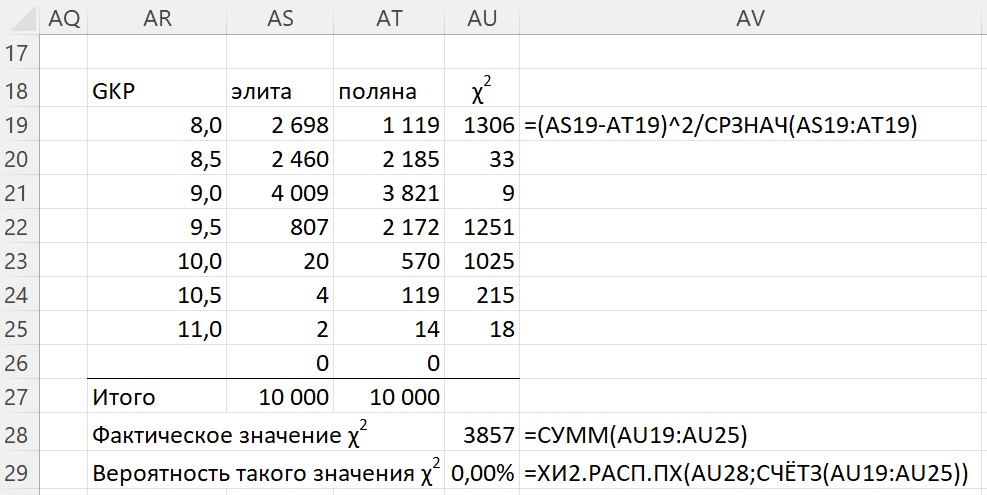


Рис. П13. Критерий χ2 для распределения двух выборок

Я проверяю значимость различий каждый раз, когда выборки кажутся похожими.